DOCUMENT RESUME

ED 481 061 TM 035 318

AUTHOR	Montgomery, Paula S.; Ranney, Lisa; Growe, Roslin
TITLE	High-Stakes Tests versus High-Quality Education.
PUB DATE	2003-10-00
NOTE	25p.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	*Educational Quality; Elementary Secondary Education; *High Stakes Tests; *State Standards; Test Coaching; Test Construction; *Test Use
IDENTIFIERS	*Louisiana; *Testing Effects

ABSTRACT

This essay discusses the uses and consequences of high-stakes tests, which are sometimes believed to interfere with high quality education. High-stakes tests put extreme pressure on students and can create a narrow curriculum with reduced instructional time because of test prepping. Using tests for high-stakes decisions causes problems because tests are not perfect, and making important decisions on limited or imperfect information can lead to bad decisions that can harm students and teachers. Strong academic standards are being developed and implemented in many states across the United States. One such state is Louisiana. The authentic standards-based reform in Louisiana is being implemented with an assessment design that aligns assessment with the content standards and benchmarks defined by the state. The Louisiana Educational Assessment Program is being developed to be a high-stakes test that reflects standards-based education directly. It is reasonable to accept that schools should be held accountable to a standards-based education in some way or other, and high-stakes tests of the sort Louisiana is developing may have a payoff in terms of a quality education. (Contains 18 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

TM TMTR 0258

ED 481 061

High-Stakes Tests Versus High -Quality Education

Paula S. Montgomery, Ph.D.

Lisa Ranney

Roslin Growe, Ed.D.

University of Louisiana at Lafayette
Lafayette, Louisiana

TM035318

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
P. Montgomery
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

BEST COPY AVAILABLE

High-Stakes Tests versus High -Quality Education

Student participation! Student knowledge! Student performance! Common educational goals! Standard-based curriculum! Authentic teaching! What creates a high-quality education? Can high-stakes testing measure whether our students are getting a high-quality education and does it reflect student achievement?

The war of assessment began over a hundred years ago. According to a schoolmaster in 1887, a teacher knew that his professional status depended upon the results he produced. He felt as though he was turned into a machine for producing results, unaccompanied by any substantial gain. Could we say that this was high-stakes testing? Definitely we can say it was for the teachers. Their salaries were depended upon the performance of their students. (Jones, 2001)

In 1966, Tyler had a vision of assessment. Part of this vision entailed individuals to be tested. They would be assigned grades, be selected to further opportunities, diagnose students to plan subsequent teaching, and evaluate the effectiveness of a curriculum or a set of teaching methods. He saw this vision develop in the National Assessment of Educational Progress (NAEP). It was designed that a national sample of students would be assessed, no student would take more than a fraction of the exercises, and no score would be obtained from any student's performance. Many of the exercises would include hands-on problems to be solved and the children would actually derive some of them. It would also represent a broad range of difficulty and educational objectives in ten different subject areas. The administrators were trained to provide high controlled assessment conditions. Furthermore, exercises were read so that it would not

harm a student because of deficiencies in reading especially in mathematics. The results of the periodic assessment would be reported, exercise by exercise, of four different age groups (9, 13, 17 years, and young adults). Concrete evidence was then available as to what the students knew and could do and how changes in performance occurred over time. It had become operational in 1969 and is still expanded today. (Jones, 2001)

Expansion and changes continue to occur in educational assessment. The exercises have changed to multiple-choice and short-answer items. The exercises have become homogeneous in difficulty. The age groups have changed to grades and the ten subject areas have received uneven attention. Mathematics, reading, science, and writing are assessed much more than literature, social studies, art, music, citizenship, and career development. Furthermore, another important difference is that test are no longer read aloud to students. (Jones, 2001)

In 1990, state-by-state assessment began in most states. Certain governmental agencies are now mandating programs of educational assessment and programs for achievement testing. However, scoring has changed since the 60s. More recently, reporting has been by achievement levels so that a comparison can be made with the actual performance to how good performance should be. Some changes may be responsible for increased pressures to report scores not just for large subpopulations but by school district, by school, by classroom, and for each individual child. (Jones, 2001)

Promotion for grades 4 and 8 and for high school diploma has been the target. This type of high-stake testing can be hazardous if a child is retained based on low-test scores. Often, these children fail to benefit from grade retention. Side effects of high-stakes testing programs are gradually being recognized. (Heubert & Hauser, 1999)

According to a teacher (2000) at Lincoln-Sudbury Regional High School in Massachusetts, it seems to be clear that these state-mandated test are designed so that at least 50% of the students have no chance to pass it. The state measures students against the new "high standards" imposed by states. These standards for some seem to impossibly high and age-inappropriate according to Gerald Bracey (2000) an education analyst. He also commented that in Fairfax County, Virginia, where students rank at the very top of international comparison in math and science, 78% still failed after intense effort. Bracey also thinks that test of this nature are very destructive educationally. These high-stakes test cover a broad range of material that teachers have to rush through the curriculum not allowing for real discussion or in-depth study. He feels as though education is reduced to memorization of disconnected facts (New Democracy Newsletter, 2000).

The crucial question to high-stakes testing is it doing more harm than good to our young people. An Associated Press story February 17, 2000 describes a little boy who hung himself with his belt from his bunk bed after leaving a note apologizing for a bad report card. He had been forced to take the standardized language test. These types of test put extreme psychological pressure on young and vulnerable people. One teacher from a vocational school in Boston recently said that in the past no more than 15% of the students had ever failed; however this year 70% of them are failing. They are convinced that they will never past the MCAS and have given up. By constantly raising the standards students have to meet, they make everyone afraid that they will never be good enough. It is subjecting our children to the same stress that employers use to control their employees on the job. (New Democracy Newsletter, 2000)

Educators argue high-stakes testing creates a narrow curriculum and reduce instruction because of test prepping. They also say that when teachers are pressured to make students better test-takers, what is lost is the rich, high-level teaching and learning that authentic, standards-based reform aims to promote. The president of the American Federation of Teachers once said, "When tests are allowed to become the be-all and end-all, they deform, not reform, education." (Thompson, 2001)

There are still many other reasons not to use any single assessment for such high-stakes consequences. Not only does it dilute curriculum and instruction, but does the assessment have high enough level of validity and reliability to justify its use as the sole basis for making consequential decisions about passing from one grade to the next. Lt. Gen. Ronald Kadish, director of the Ballistic Missile Defense Organization once said, "I don't think we should draw conclusions from any one test that are irrevocable. No one test tells you everything you need to know." Another problem that exists is that the tests are frequently misused. They are often used to evaluate teachers and schools. (Thompson 2001)

It seems as though that test that have no specific decision tied to them can become high stakes to teachers and administrators when they must face public pressure after the scores are made public. Also, low-stakes test can be transformed into high-stakes test at a school district level when educational decisions are based on the test results. (Reading Teacher, 1999)

Are our children receiving the quality of education needed for successful life and work in our rapidly changing world? Isn't the right thing to do is to provide a high-quality education for all children and provide them with lifelong consequences? Then

why are we focussing all of our energy and resources at trying to improve learning and student performance that is measured by a single assessment? The question arises then is it possible to require all students to meet a set of rigorous standards in order to graduate from high school without using only a single test as the means of determining whether those standards have been met. (Gratz, 2000)

High school graduation and passing scores continue to increase pressure on students and teachers. Students from schools in affluent neighborhoods pass these exams; however, many from socioculturally disadvantaged districts do not. Unless these students are provided ample assistance, they cannot perform well. (Verones, 2000)

Are our students with disabilities given accommodations on statewide assessment tests? According to the Disabilities Education Act (IDEA) Amendments of 1997 it was mandated the students with disabilities participate in large-scale assessment programs. IDEA mandates the use of accommodations to maximize their participation and to ensure that the results lead to valid decisions. It is the local and state educational agencies that provide policies to guide the decision of who participates and what types of accommodations and modifications are provided and not IDEA. Many states want to use the Individualized Education Program (IEP) teams to make these decisions, which has caused some uncertainty. Who ever makes that decisions, they must make critical decisions about whether students can take a test in a standard fashion. They need to make decisions whether they need accommodations, and if so what kind. Remember accommodations do not change the nature of the construct being tested, but can affect a student's or group's performance in comparison to a peer group. Accommodations could be as simply providing time for certain students or groups of students to complete the test

and tasks without making any other accommodations. Whereas, modifications result in a change in the test of how it is given, how it is completed, or what construct is being assessed. Unfortunately, because many teachers are not as knowledgeable about the influence of disabilities on student performance on test performance, students are not receiving the accommodations they need. Wise (1991) reported that most teacher certification agencies at the state level do not require assessment/measurement courses for initial certification. Schafer (1991) discovered that "only about half of the teacher education programs in the nation require a course in measurement for initial certification." To guarantee stability and consistency across districts and teachers, the Department of Education for each state, may want to set guidelines for appropriate modifications and accommodations used for testing. (Hollenbeck & Tindal, 1998)

Important decisions are made from high-stakes testing other than just promoting from one grade to another. If students score high on a single test they could be placed in honors classes or a gifted program. However, if a student scores low on a high-stakes test, it could mean that they would be placed in low-level class or they could even be rejected to a particular college. The Board of Directors of the International Reading Association is opposed to high-stakes testing in which single test scores are used to make important educational decisions. Their central concern is that they feel as though testing has become a means of controlling instruction as opposed to a way of gathering information to help students become better readers. (Reading Teacher, 1999)

Why does using test for high-stakes decision cause problems? Tests are not perfect and making important decisions on limited or imperfect information can lead to bad decisions that can harm the students and teachers. Another problem is the tendency

to narrow the curriculum and inflate the importance of the test. As we all know, schools should address a broad range of student learning needs, and not just subjects or parts of subjects. Because of the pressure that teachers face with high-stakes test, they will tend to focus their efforts on those activities that they think will improve the single important score. Fine arts, physical education, and the sciences are being neglected because so much time is spent focussing on reading, writing, and mathematics. In 1999 a survey was taken by Hoffman who revealed in one state that uses high-stakes assessments that 75% of classroom teachers surveyed thought the state assessment had a negative impact on their teaching. Furthermore, teachers sometimes respond to test pressure by focussing their attention on particular students. This means that only low-performing students may get attention, where as others are ignored. Another negative result of high-stakes tests is the loss of instructional time that is needed to prepare for the testing. (Reading Teacher, 1999)

With all the controversy on high-stakes tests, then why have this type of testing system? The education reform movements in recent years have focused on developing new standards and assessments for students. Actually, 47 states have adopted new standards for student learning and most have adopted statewide testing systems. This standard-based reform envisioned that high-stakes tests would promote student learning. It would promote it by implementing a high quality curriculum framework and assessments would be tied to these standards. Secondly, it would promote course offerings that reflect this high quality curriculum. Furthermore, teacher preparation and professional development would be guided by related standards for teaching. Lastly, that student testing would address the quality of teaching, the allocation of resources, or the

nature of schooling. The major problem of reform is seen as a lack of effort on the part of educators and students. It is believed that standards will motivate change if they are used to apply consequences to those who fail to meet them. The use of scores on achievement tests is used to make decisions that have important consequences for the students and teachers to promote accountability. (Darling, 2002)

Does high-stakes test have a favorable impact on student achievement? Does a curriculum-based exit exam system improve the teaching and learning of core subjects? According to the Competitiveness Policy Council, that "external assessments be given to individual students at the secondary level and that the results should be a major but not exclusive factor qualifying for college and better jobs at better wages." (1998)

The following are characteristics that define a curriculum-based external exit examination system (CBEEESS):

1. It produces signals of student accomplishment that have real consequences for the student.
2. It defines achievement relative to a standard, not relative to other students in the classroom or school.
3. It is organized by discipline and keyed to the content of specific course sequences.
4. It signals multiple levels of achievement in the subject.
5. It covers almost all secondary school students. (Bishop, 1998)

A minimum competency exam that requires students to pass to graduate in secondary schools is not a CBEEES because they fail characteristics three and four. These test focus on basic skills taught in primary and the lower grades of secondary

schools and the passing standard is quite low. The exit exam for high school is generally taken in 9th or 10th grade, and most students pass it the first time. The high school transcripts only indicate whether the student eventually passed the exam, not achievement levels above the minimum. Also, the majority of the students who pass the exam for the first time are not stimulated for further study. (Bishop, 1998)

According to Thompson (2001) we should be interested in students who can produce high-quality work rather than students who have mastered the ability to take standardized test. In the long run these are the people who will be rewarded in their personal and professional lives after graduation, when test-taking skills are no longer relevant.

Since high-stakes test seems to be a trend in this decade and seems to continue to be part of the educational system, it is important that we have a quality assessment plan if the results are to have such a major long-term consequence. Testing students' knowledge and skills is definitely an important part of education; however, it is not the only type of educational assessment. The main purpose of assessment is to help students by providing information about how instruction can be improved. The International Reading Association has made recommendations:

- The teachers should create rich assessment environments in their classrooms and schools. They should construct systematic and rigorous assessments for students that will help them gain confidence with the standards. Teachers should also take responsibility to educate parents, community members, and policy makers about he forms of classroom-based assessment used in addition to standardized tests. They should also understand the difference between ethical and unethical practices when

teaching to the test. It is ethical to get students familiar with the format of the test so they are familiar with the types of questions and responses required. It would not be ethical to devote more time teaching the test than time that would normally be used for regular instruction.

- Researchers should continue to investigate how assessment can better serve our educational goals. Evaluations of high-stakes tests should be ongoing process. The results should include the impact of the curriculum, time in testing and test preparation, the costs of the test, parent and community communication, and effects on teacher and student motivations. Good baseline data and follow-up studies should be conducted.
- Parents and the community members should help in bringing a balance to the assessment design. They should be able to ask questions how tests are impacting their children, cost, time, and alternative methods. They should lobby for the development of classroom-based forms of assessment that provide useful, understandable information and improve instruction.
- Policy makers should design an assessment plan that considers the complexity of reading, learning to read, and the teaching of reading. Multiple measures should be assessed rather than just performance on a single test. England has adopted an assessment system that focuses on teacher informal assessments, ongoing performance assessments, portfolios, teacher recommendations, and standardized testing. These types of sources provide more valid decision-making. The International Reading Association also recommends sampling strategies to encourage careful inspection of issues of validity and reliability. They also suggest not using

incentives, money, or recognition of test scores to reward or punishing schools or teachers. Lastly, policy makers should visit the classroom, listen to teachers talk about the curriculum and decisions they are making, and talk to teachers about the kinds of assessments they use in the classroom. (Reading Teacher, 1999)

Millions of students this past year were judged by some type of test scores which are keyed to state standards. The question to ask is there a problem with the state standards or the high-stakes tests or both? The Thomas B. Fordham Foundation (1998) researched to find out the answers to such questions as: Are the state standards rigorous? Clear? Could a teacher, parent, or student pick them up and make sense of them? Are they likely to boost student achievement? Do they lend themselves to assessment and accountability? Do they point schools in the right direction? Their studies showed that some states did well in some subjects; however, very few did well overall. The main conclusion was that most states had a very long way to go before their academic standards would be strong enough to bear the burden of high-stakes tests. They compiled a grade point average (using a traditional four-point scale) for each state of a grade of A, B, C, D, and F. No state received an A, and only three states, California, Texas, and Arizona received a grade of a B. Nine states flunked and the national cumulative GPA was 1.3. (Finn, Petrill, & Vanourek, 1998)

Why such a low marks on state academic standards by reviewers of The Thomas B. Fordham Foundation (1998)? These reviewers identified four problems.

1. Many state standards are extremely vague. Academic standards must be clear, specific, and measurable. Unfortunately, many states have produced very vague documents and the states supply many excuses for their vagueness.

2. Many state standards are hostile to knowledge. Flinn, Petrilli, and Vanourek (1998) searched for specific knowledge and skills that the states want students to master; however, in many of the classrooms found an acute shortage.
3. Many state standards are entranced by "relevance." An English-standards reviewer Sandra Stotsky (1998) once said, "To require students at higher educational levels to read their lives into the literature they are asked to study undermines the very capacity of a literary work to help readers transcend their limited experiences. A major function of literature is to expand perspectives and free students from insularity." Education should widen a student's horizons. Standard writers should be interested in identifying important knowledge, and not always to ensure that a student will be immediately gratified. Good teachers will find ways to make material come alive for students.
4. Many states wrote standards of teaching rather than standards of learning. The educational standards should be clear as to what the student will learn at the different grade levels and how well this is to be learned. Teaching methods, classroom strategies, and lesson plans should be up to the teacher, and not up to the state. (Finn, Petrilli, & Vanourek, 1998)

Standards should help schools focus on results, and this will enable school models to emerge so that a range of choices can better serve the needs and learning styles of children and the passions and talents of teachers. It is mentioned in this Fordham report by Susan Traiman (1998) that, "We must not get endlessly stuck in the process of perfecting standards, a process that dooms standards-based reform to a state of paralysis.

But let's put it this way: In standards jargon, most of the documents we currently have are below basic. Let's at least get them up to proficient." (Finn, Petrilli, and Vanourek 1998)

The United States for some time has lacked meaningful standards and avoided real accountability. Times are starting to change. This change relies on standards, assessments, and consequences. The standards describe what students should know and be able to do in core subjects at critical points in their education. Standards should define the desired results of schooling - what students know and be able to do by the time they graduate from high school. With the accountability system that most states have implemented, these results are judged by the students' achievements. However, such a system depends on strong academic standards. (Flinn & Petrilli, 2000)

Strong academic standards are being developed and implemented in many states across the United States. One state in particular that is starting to succeed in providing a high-quality education to students is Louisiana. Their authentic standards-based reform involves teachers, students, parents and others as active participants in developing and refining common learning standards. This reform did not take place overnight - it has taken the past decade to develop.

One concern that was presented by the Competitiveness Policy Council (1998) concerning high stakes tests was that it didn't address multiple levels of achievement. Most states just have a pass or fail system. Under the Louisiana Educational Assessment Program (2000-2001) the students will no longer receive a simple "pass/fail" score; they will receive one of the following five achievement levels.

- Advanced: A student at this level has demonstrated superior performance beyond the proficient level of mastery.

- Proficient: A student at this level has demonstrated competency over challenging subject matter and is well prepared for the next level of schooling.
- Basic: A student at this level has demonstrated only the fundamental knowledge and skills needed for the next level of schooling.
- Approaching Basic: A student at this level has only partially demonstrated the fundamental knowledge and skills needed for the next level of schooling.
- Unsatisfactory: A student at this level has not demonstrated the fundamental knowledge and skills needed for the next level of schooling. (Louisiana Assessment Program 2000-2001)

Another suggestion by the Competitiveness Policy Council (1998) that exams (high-stakes) be organized by discipline and keyed to content. The Louisiana Department of Education has been involved with efforts for reform since 1993. However, in May 1997 the State Board of Elementary and Secondary Education approved content standards in English language arts, mathematics, science, and social studies, foreign language and the arts. These standards reflect the essential knowledge and skills that students need to become good scholars and productive citizens. The following skills provide a base for all of the content standards to make learning more meaningful.

1) Communication
2) Problem Solving
3) Resource Access and Utilization
4) Linking and generating knowledge
5) Citizenship

The Louisiana Department of Education in (1997) had also initiated new criterion-referenced test to align with the content standards in four of the six areas: English language arts, mathematics, science, and social studies. The criterion-reference tests were to be administered in grades 4, 8, and 10. These grades are consistent with the grades at which the content standards and benchmarks are clustered (k-4, 5-8, and 9-12), as well as with grades assessed by the National Assessment of Educational Progress (NAEP). The remaining state criterion-referenced portions of the GEE 21 (Graduate Exit Exam) at grade 11 was to be implemented in the Spring 2002. (Louisiana Assessment Program 2000-2001)

An issue discussed by Lt.Gen. Ronald Kadish (1999) was that assessments don't seem to have a high enough level of validity and reliability. This issue was addressed by the Louisiana Department of Education by first developing an assessment design that would align the assessment with the content standards and benchmarks. Then testing contractors developed test items using the assessment specifications. Then an additional review was conducted with a Bias Review Committee that viewed the items for sensitive or biased material regarding gender, ethnicity, and issues related to special populations of students. The revised items were then submitted for final approval and then prepared for field-testing. The fourth and eighth grades were initially field tested in Mathematics and English Language Arts and then in the Spring of 1999 Science and Social Studies were field-tested. English Language Arts and Mathematics items in 10^{th} grade were first field tested in Spring 2000. The schools that participated in the field test were chosen randomly based on the state's school sub-populations on the factors of ethnicity, socioeconomic status, school size, and school achievement performance. The data

collected were then submitted to the Assessment Advisory Committees and the Bias Review Committee for a final review. Finally the Division of Student Standards and Assessments in collaboration with the testing contractors assembled the initial LEAP 21 (Louisiana Educational Assessment Program) tests in English Language Arts and Mathematics for Grade 4 and Grade 8. These high-stakes tests were implemented in March 1999. The same procedures were followed to create the LEAP 21 for Science and Social Studies. Science and Social Studies was implemented into the high-stakes test in March 2000. In March 2001 the GEE 21 in Grade 10 for English Language Arts and Mathematics tests were added. The GEE 21 for Grade 10 for Science and Social Studies was to be added and implemented in Spring 2002. (Louisiana Assessment Program 2000-2001)

The LEAP 21 carries high-stakes for those students in Grade 4 and Grade 8. They must at least reach the Approaching Basic Level to be promoted to the next grade level. The GEE 21 for Grades 9 - 12 must reach the Approaching Basic Level to be eligible for a high school diploma. Since these tests carry such high stakes for students, the state of Louisiana has devised intensive summer remediation only for those students who receive an unsatisfactory score. These students would then be retested at the end of the summer. (Louisiana Assessment Program 2000-2001)

Furthermore, to yield valid and reliable longitudinal data, the difficulty level of the test must be equivalent from year to year. A process called "test equating" is used to maintain consistency by scaling the scores. This allows the use of raw scores to compute students' scaled scores and to establish a common achievement level standard from test to test. (Louisiana Assessment Program 2000-2001)

The scaled score ranges between 100 and 500 for all grades and content areas for the LEAP 21 and GEE 21. The following tables show the scaled score range for each of the five levels.

Achievement Level	English Language Arts Scaled Score Range		
	Grade 4	Grade 8	Grade 10
Advanced	408-500	402-500	398-500
Proficient	354-407	356-401	347-397
Basic	301-353	315-355	299-346
Approaching Basic	263-300	269-314	270-298
Unsatisfactory	100-262	100-269	100-281

Achievement Level	Mathematics Scaled Score Range		
	Grade 4	Grade 8	Grade 10
Advanced	419-500	398-500	377-500
Proficient	370-418	376-397	346-376
Basic	315-369	321-375	305-345
Approaching Basic	282-314	296-320	286-304
Unsatisfactory	100-281	100-295	100-285

Achievement Level	Science Scaled Score Range	
	Grade 4	Grade 8
Advanced	405-500	400-500
Proficient	360-404	345-399
Basic	306-359	305-344
Approaching Basic	263-305	267-304
Unsatisfactory	100-262	100-266

Achievement Level	Social Studies Scaled Score Range	
	Grade 4	Grade 8
Advanced	399-500	404-500
Proficient	353-398	350-403
Basic	301-352	297-349
Approaching Basic	272-300	263-296
Unsatisfactory	100-271	100-262

The English Language Arts test for each grade has four parts; however, competence in speaking and listening is not currently incorporated in LEAP 21 and GEE 21. The Louisiana State Department of Education is presently exploring ways to incorporate this into the standard at the local level.

The parts included are:

1. Writing: It requires students to produce a composition in response to a writing topic.
2. Using Information Resources: It requires students to complete a specified task designed to measure how a student can locate, select, and synthesize information from a variety of texts, media, references, and technological sources to acquire and communicate knowledge.
3. Reading and Responding: It includes four reading passages and a variety of item types. It includes multiple-choice and short-answer items. Grades 8 and 10 include an essay question that requires students to comprehend and react to the content of the reading passages.
4. Proofreading: Requires students to read a text that includes mistakes in grammar, usage, and mechanics. It also requires students to answer multiple-choice questions that require choosing the best way to correct each mistake. (Louisiana Assessment Program 2000-2001)

The Mathematics tests are given to Grades 4, 8, and 10 and consist of two major parts.

Part 1: Consist of concepts and skills in all six strands of mathematics in multiple-choice format. The six strands include:

1. Number Relations: Problem solving investigations, students demonstrate an understanding of the real number system and communicate the relationships within that system using a variety of techniques and tools.
2. Algebra Standard: Students demonstrate an understanding of concepts and processes that allow them to analyze, represent, and describe relationships among variable quantities and to apply algebraic methods to real-world situations.
3. Measurement Standards: Students demonstrate an understanding of the concepts, processes, and real-life applications of measurement.
4. Geometry Standard: Students demonstrate an understanding of geometric concepts and applications involving one-, two-, and three-dimensional geometry and justify their findings.
5. Data Analysis, Probability, and Discrete Math Standard: Students discover trends, formulate conjectures regarding cause-and-effect relationships, and demonstrate critical-thinking skills in order to make informed decisions.
6. Patterns, Relations, and Functions Standard: Students demonstrate an understanding of patterns, relations, and functions that represent and explain real-world situations.

Part 2: Consists of four complex mathematical tasks at grades 8 and 10 and three for grade 4. All of these problems involve a number of separate steps and require application of multiple skills. These tasks are open-ended and include numerical answers, short written answers, and other types of constructed responses. The response by the student is scored analytically for such traits as accuracy of the

answer, proper operations used, and appropriate problem-solving approach or strategy. Calculators are permitted on Part 2 and partial credit is given. (Louisiana Assessment Program 2000-2001)

The Science portion of this high-stakes tests require students to use their content knowledge of science to explain, connect, and apply contents to new situations. The Science Test consists of three parts:

Part 1: The first part consists of 40 multiple-choice questions that assess concepts and skills in all five strands of science. The five strands are:

1. Science as Inquiry
2. Physical Science
3. Life Science
4. Earth and Space Science
5. Science and the Environment

Part 2: This part consist of four short-answer questions that assess four content strands: Physical Science, Life Science, Earth and Space Science, and Science and the Environment. The questions on this part of the test will allow students to reflect on an idea, demonstrate their understanding of concepts and the processes of science, making meaning of a given set of data, and critique the information.

Part 3: This part consists of a comprehensive science task that requires Grade 4 students to observe, utilize, and react to materials in an investigation and draw conclusions based on their experiences. Grade 8 students will respond to a scenario that requires scientific investigation. At both grades the task/scenario integrates the Science as Inquiry content strand with at least one other content strand. The

questions vary in format: constructed response, data tables, short answer, and at least one essay question. (Louisiana Assessment Program 2000-2001)

The Social Studies tests challenge students to expand their thinking in Social Studies and become accomplished problem solvers and informed decision-makers. It consist of two major parts:

Part 1: The first part of the test consists of fifty multiple-choice test items for Grade 4 and sixty for Grade 8. It assesses knowledge, conceptual understanding, and application of skills in all four social studies strands: Geography, Civics, Economics, and History. These strands are intermingled and not arranged into separate sections.

Part 2: This part consists of four open-ended questions or tasks that requires a constructed response and requiring higher-order thinking in social studies. The student may have to construct or interpret a chart, graph, map, timeline, or other graphic representation. (Louisiana Assessment Program 2000-2001)

High-stakes tests of this nature that directly reflect standards-based education reform, perhaps may have a considerable payoff in terms of a high-quality education. It is reasonable to accept that schools should be held accountable to a standards-based education in some way or another. Parents and teachers need to understand that effective education must foster curiosity and a child's desire to learn, even if at the expense of high-stakes tests. William Butler Yeats (2001) once said, "Education is not the filling of a pail, but the lighting of a fire."

References

Bishop, J. (Spring 1998). The effect of curriculum-based external exam systems. Journal of Economic Education. Vol.29, Issue 2, p.171-183. Retrieved June 12, 2002 from http://ftviewer.epnet.com/ftviewer/amfulltext.asp?EBSCOcontenthtm

Darling, L. (Jan 2002). What's at stake in high-stakes testing? The Brown University Child and Adolescent Behavior Letter. Retrieved June 13, 2002 from http://www.Findarticles.com/cf_dls/m0537/1-83139498/pl/article.jhtm

Downing, S. & Haladyna T. (1996). A model for evaluation of high-stakes testing programs: Why the fox should not guard the chicken coop. Educational Measurement: Issues and Practice. Vol. 5, Issue 1, p.5-12.

Finn, C. E. Jr. & Petrilli, M. J. (2000). The state of state standards. The Thomas B. Fordham Foundation Report. Retrieved June 16, 2002 from http://www.edexcellence.net/topics/standards.htm.

Finn, C. E. Jr., Petrilli, M., & Vanourek, G. (1999). The state of state standards: Four reasons why most don't cut the mustard. Education Week. November 11, 1998. Retrieved June 16, 2002 from http://www.edexcellence.net/library/mustard.htm

Firestone, W. A. & Mayrowetz, D. (Aug 2000). Rethinking high stakes: Lessons from the U.S. and England and Wales. Vol. 102, Issue 4, p.724-750. Retrieved June 12, 2002 from http://ftviewer.epnet.com/ftviewer/amfulltext.asp?EBSCOcontenthtm

Gratz, D. B. (May 2000). High standards for whom? Phi Delta Kappan, p.684.

Heuber, J.P. & Hauser, R.M. (1999). High stakes: Testing for tracking promotion and graduation. Washington, D.C.: National Academy Press. Retrieved June 12, 2002 from http://ftviewer.epnet.com/ftviewerdelivery.asp?deliveryoptionhtm

Hollenbeck, K. & Tindal, G. (Fall 98). Teachers' knowledge of accommodations as validity issue in high-stakes testing. Journal of Special Education. Vol. 32, Issue 3, p. 175-184. Retrieved June 12, 2002 from http://ftviewer.epnet.com/ftviewer/amfulltext.asp?EBSCOcontenthtm

International Reading Association (Nov 1999). High-stakes assessments in reading. Reading Teacher, Vol.53, Issue 3, p.257-264. Retrieved June 12, 2002 from http://ftviewer.epnet.com/ftvieweramfulltest.asp?EBSCOcontenthtm

Jones, L. V. (Feb 2001). Assessing achievement versus high-stakes testing: A crucial contrast. Educational Assessment. Vol. 7, Issue 1, p.21-29. Retrieved June 12, 2002 from http://ftviewer.epnet.com/ftviewer/delivery.asp?deliveryoptionhtm

LEAP 21/GEE 21 (2000 -2001). Annual Report Section 1: The Louisiana Educational Assessment Program. Retrieved June 13, 2002 from http://doe.state.la.us/DOE/account//leap.asphtm

New Democracy Newsletter (May-June 2000). Why are they doing this to our kids? Retrieved June 13, 2002 from http://www.newdemocracyworld.org/testing.htm

Reising, B. (Sep/Oct 2000). High School Exit Exams. Clearing House, Vol. 74, Issue 1, p.4-5. Retrieved June 12, 2002 from http://ftviewer.epnet.com/ftviewer/amfulltext.asp?EBSCOcontenthtm

Schafer, W.D. (1991). Essential assessment skills in professional education of teachers. Educational Measurement Issues and Practice. Vol. 10, Issue 1, p.3-6)

Thompson, S. (Jan 2001). The authentic standards movement and its evil twin. Phi Delta Kappan. Vol.82, Issue 5, p.358-363. Retrieved June 12, 2002 from http://ftviewer.epnet.com/ftviewer/delivery.asp?deliveryoptionhtm

Veronesi, P. (Sept/Oct 2000). Testing and assessment in science education: Looking past the scoreboard. Journal of Teacher Education. Vol. 74, Issue 1, p.27-31. Retrieved June 12, 2002 from http://ftviewer.epnet.com/ftviewer/deliver.asp?deliveryoptionhtm

Wise, S.L., Lukin, L.E. & Ross, L.L. (1991). Teacher beliefs about training in testing and measurement. Journal of Teacher Education, Vol. 42, Issue 1, p.37-42.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

TM035318

I. DOCUMENT IDENTIFICATION:

Title: High-Stakes Tests Versus High-Quality Education

Author(s): Paula S. Montgomery, Ph.D., Lisa Ranney, Roslin Grove, Ed.D.

Corporate Source:

Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☒

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: Paula S. Montgomery / Roslin Grove / Lisa Ranney

Printed Name/Position/Title:

Organization/Address: University of Louisiana at Lafayette, 1217 Olde Street New Iberia, LA 70560

Telephone: 337-482-6745 FAX: 482-5262

E-Mail Address: pom8849@louisiana.edu Date: 10/13/03

Paula S. Montgomery, Ph.D. Educational Foundations and Leadership Assistant Professor
Lisa Ranney Parochial School Teacher
Roslin Grove, Ed.D. Educational Foundations and Leadership Department Chair/Professor

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfacility.org